

## 4.11.2 Scale Factor Studies

j. guimarães da costa, a. holloway, d. sherman, s. rappoccio

effect of recent cuts

effect of slight changes to method

effect of gross changes to method

# Samples

We use the same initial datasets as TomW and HenriB:

- SecVtx from this summer (blessed)
- 4.11.2 reprocessing → TopNtuples
- “Top/EWK Version 4” good runs: Ele, Silicon, Muon

⇒ Not identical (crashed jobs...)

Sample	dataset	N(events) before cuts
New Electron MC	btop2a	507779
8 GeV Electron Data	b2t120,b2t140	1799511

# Event/Object Selection : Common cuts

## *∴ Electron selection*

- $E_T > 9.0 \text{ GeV}$
- $0.5 < \frac{E}{P} < 2.0, \frac{\text{had}}{\text{em}} < 0.05$
- $L_{\text{shr}} < 0.2$
- $|\Delta x_{\text{CES}}| < 3 \text{ cm}, |\Delta z_{\text{CES}}| < 5 \text{ cm}, \chi_{\text{CES strip}}^2 < 10$
- $|z_e - z_0| < 5 \text{ cm}$
- $\text{Isol} > 0.1$

## *∴ Object selection*

- Highest  $E_T$  electron
- Nearest e-jet with  $\delta R_{\text{e-jet,e}} < 0.4$
- $\delta\phi_{\text{e-jet,a-jet}} > 2.0$
- $|\eta_{\text{a-jet}}| < 1.5$
- Conversions:  $\delta \cot(\theta) < 0.04, d_{xy} < 0.2 \text{ cm}$

## Event/Object Selection : Different cuts

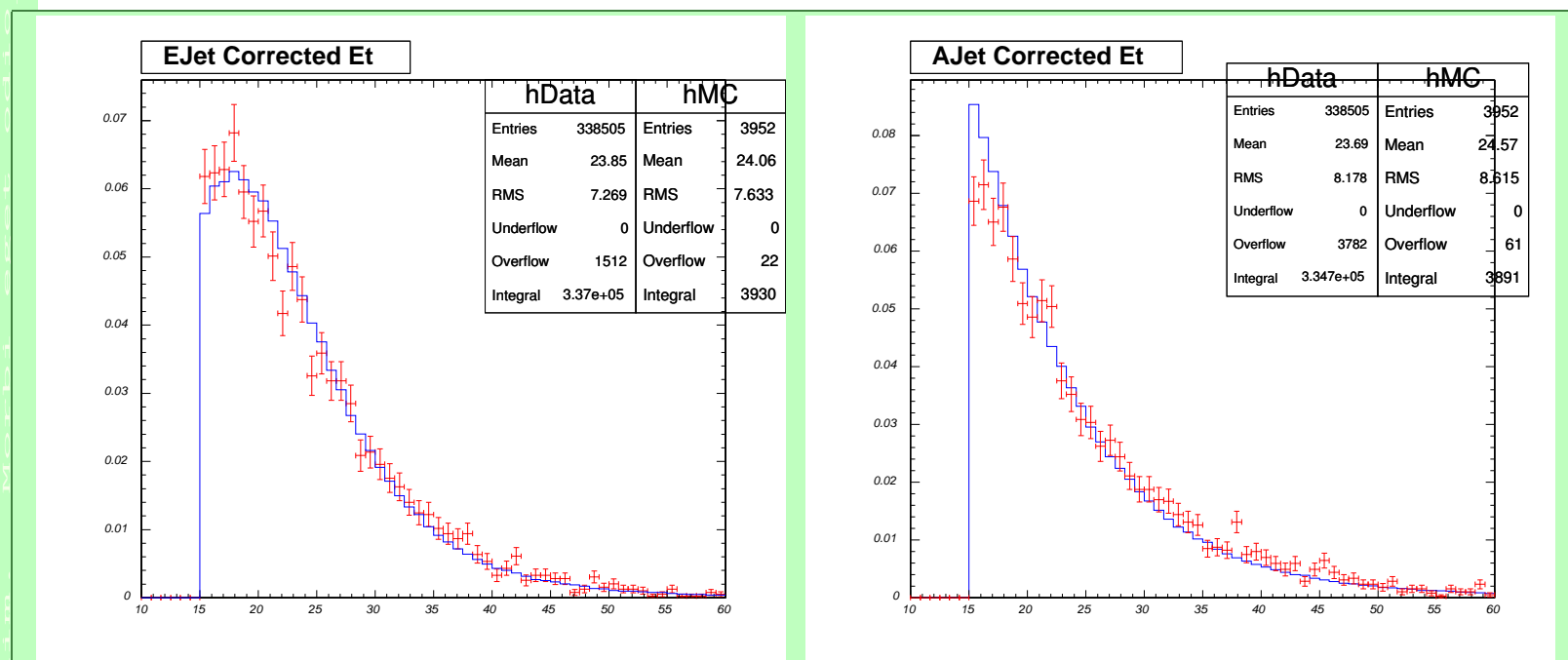
Quantity	Blessed+Trigger	HenriB	TomW (Jan16)
Electron $p_T$	4.5 GeV	8.0 GeV	9.0 GeV ( $E_T > 12$ )
Electron track	—	SVX fiducial	—
E-jet $E_T$	10 GeV	15 GeV (Level4)	10 GeV
A-jet $E_T$	12 GeV	15 GeV (Level4)	12 GeV
Which A-jet?	Highest $\delta_\phi$	Highest $E_T$	Highest $\delta_\phi$

With these three sets of cuts:

- we compared distributions in data and simulated events
- we measured a ( $SF$ ) to check consistency

# Data/MC Comparisons

“HenriB” cuts with the trigger simulation:



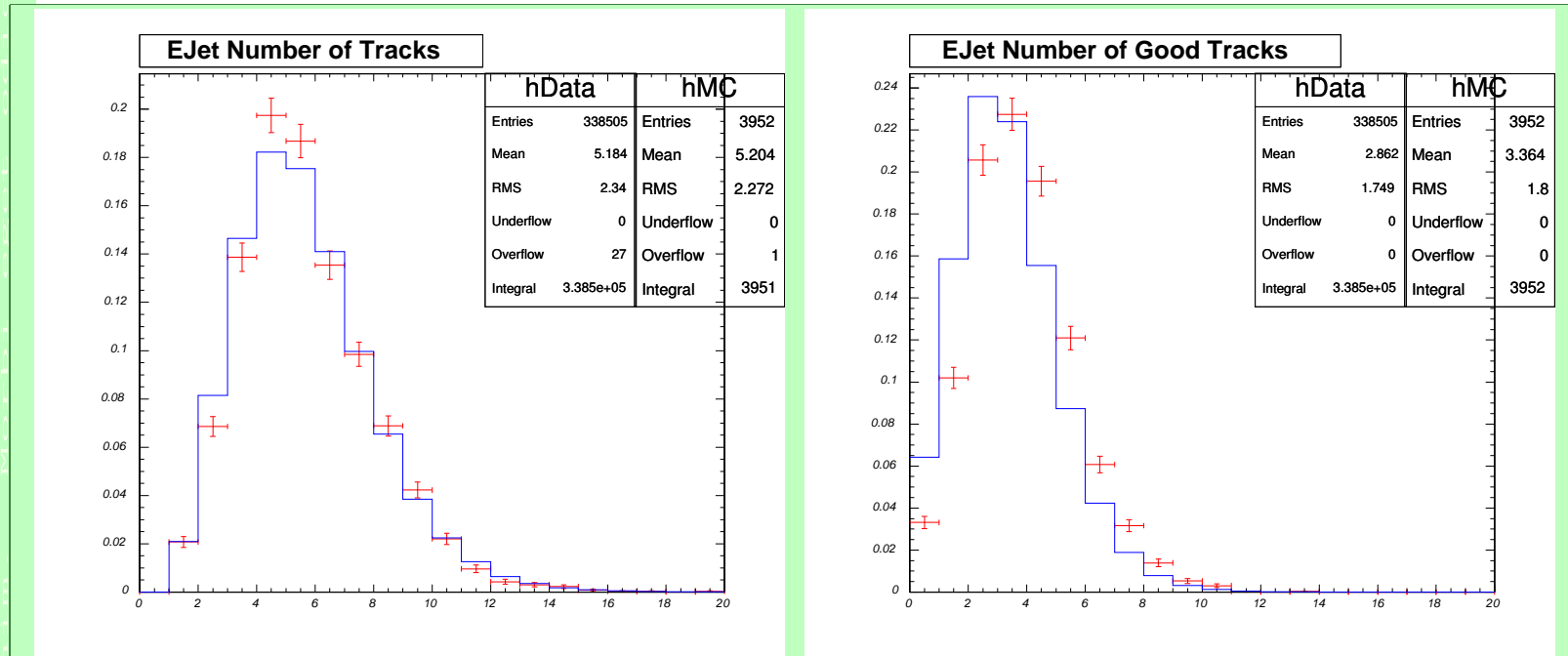
but how much agreement do we expect?

- Untagged distributions:  $\boxed{F_{MC}}^H \sim 0.85$  but  $\boxed{F}^H \sim 0.2$
- Electron characteristics:  $\frac{\boxed{N_{conv}}}{\boxed{N}} + \boxed{F}^H \sim 0.5 \Rightarrow$  electrons from ( $\pi^\pm$  fakes, Drell-Yan...)

*The important thing for a useful ( $SF$ ) is to get  $\frac{\boxed{F}^b}{\boxed{F}^H}$  correct.*

## Data/MC Comparisons: continued

An important thing for a  $(SF) \sim 1$  is to get silicon hits correct:



- $N_{\text{trk}}$  agrees better than  $N_{\text{good}}$

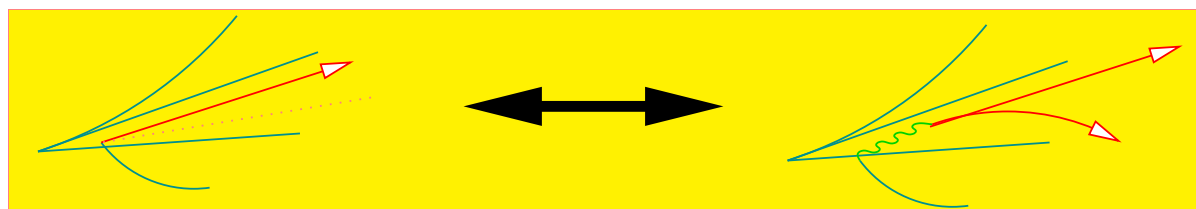
## Changes to $(SF)$ Measurement

The well-known blessed **efficiency** formula (condensed):

$$\delta^H \equiv (\epsilon^{+H} - \epsilon^{+L}) \simeq \frac{[N]_+^+ - [N]_+^- - [N]_-^+ + [N]_-^-}{([N]_+ - [N]_-) - (1 - [F]^H) [N](X)}$$

- We're neglecting the  $L_{xy}$  asymmetry in light-flavor tags for now.
- I'm hiding the fraction of light flavor electron jets that have tagged, heavy flavor away jets under " $X$ ." This is what we measure in Run II with conversions.
- $X$  has changed to account for **conversions in HF jets** – the details will be documented<sup>1</sup>.

**The tradeoff:** now conversion finding and tagging are assumed independent.



<sup>1</sup>We use  $X = \frac{(C^+ - C^-)([N]_+ - [N]_-) - (C_+ - C_-)([N]_+^+ - [N]_-^-)}{(C^+ - C^-)N - C([N]_+^+ - [N]_-^-)}$  where  $C_j^i \equiv ([N]_{\text{conv}}^i)_j - \epsilon^0([N]_j^i)$ : in the summer,  $(C^+ - C^-) \rightarrow 0$ .

## Mostly Summer Method

Measuring the efficiency in order to get  $(SF)$  has drawbacks:

- It depends on  $[F]^H$  (measured with  $D^0$  reconstruction).

Instead, let's eliminate  $[F]^H$  using  $(SF)$ .

For interested parties:

$$[F]^H = \frac{([N]^+ - [N]^-)}{[N]\delta^H} = \frac{([N]^+ - [N]^-)}{[N] \left( (SF) \times \delta_{MC;ST}^H \right)}$$

The result<sup>2</sup> is a measurement of  $(SF)$  – you have to multiply by  $\delta_{MC;DT}^H$  from “truth” info if you need the double-tag efficiency.

<sup>2</sup>The next steps are simple substitution:

$$\delta^H = \frac{[N]_+^+ - [N]_+^- - [N]_-^+ + [N]_-^-}{[N]_+ - [N]_- + [N] \left( 1 - \frac{([N]^+ - [N]^-)}{[N]((SF) \times \delta_{MC;ST}^H)} \right) (X)},$$

so

$$(SF) = \frac{\frac{([N]_+^+ - [N]_+^- - [N]_-^+ + [N]_-^-)}{\delta_{MC;DT}^H} - (X) \frac{([N]^+ - [N]^-)}{\delta_{MC;ST}^H}}{[N]_+ - [N]_- - [N](X)}$$

It's important to distinguish the MC single and double tag efficiencies  $\delta_{MC;ST}^H$  and  $\delta_{MC;DT}^H$  – they aren't the same.



## Results: Mostly Summer Method

So, using the same cuts but different algebra:

Talk	$\overline{F}^H$	$\delta^H$	$(SF)$ (stat errors)
<i>HenriB (01/30)</i>	<i>0.223</i>	<i>25.1%</i>	<i>.854<math>\pm</math>.020<math>\pm</math>0.058</i>
<i>TomW (01/30)</i>	<i>0.291</i>	<i>22.4%</i>	<i>.892<math>\pm</math>.020<math>\pm</math>0.054</i>
blessed cuts(02/13)	0.241	21.3%	.766 $\pm$ .021 $\pm$ .070
HB cuts(02/13)	0.230	23.5%	.784 $\pm$ .023 $\pm$ .082
TW-tight cuts(02/13)	0.209	22.7%	.835 $\pm$ .030 $\pm$ .104

We have a little more data (?) but a lot less MC – MC statistical errors dominate.

## Fit-based Method

- We can't escape the assumption that **e-jet tag efficiency is completely uncorrelated with a-jet characteristics**, but we no longer assume
  - that the efficiency to tag the e-jet and the efficiency to find a conversion in it are uncorrelated
  - that Run 151435 represents the geometric correlations of double-tags accurately, or
  - that light flavor tags are (equal/proportional) to negative  $L_{xy}$  tags.

This leads to the **simple** model:

$$\frac{N_{\text{sample}}^+}{N_{\text{sample}}} = (\epsilon^+)^L F_{\text{sample}}^L + (\epsilon^+)^H F_{\text{sample}}^H$$
$$\frac{N_{\text{sample}}^-}{N_{\text{sample}}} = (\epsilon^-)^L F_{\text{sample}}^L + (\epsilon^-)^H F_{\text{sample}}^H$$
$$1 = F_{\text{sample}}^L + F_{\text{sample}}^H$$

## Results: New Method

We fit to this model with many subsamples of the data, divided based on the away-jet characteristics:

no 7 GeV electron, $N_{\text{trk}} \leq 5$	no 7 GeV electron, $N_{\text{trk}} > 5$
7 GeV electron, $N_{\text{trk}} \leq 5$	7 GeV electron, $N_{\text{trk}} > 5$

We're trading statistical power for a grasp on systematics.

	$\epsilon^+$	$\epsilon^-$	$\boxed{F}^H$
Data:	$21.55 \pm 0.03\%$	$0.89 \pm 0.07\%$	$.218 \pm .009\%$
MC:	$24.75 \pm 8.16$	$0.25 \pm 0.42\%$	$.84 \pm .15$
(truth):	$24.62 \pm 0.73$	$0.56 \pm 0.11\%$	$.858 \pm .012$

With fit values,  $(SF) = 0.84$  with large errors from MC statistics.  
Using truth,  $(SF) = 0.859 \pm 0.055$

## Conclusions

- Hints that  $(SF)$  varies with electron  $p_T$  cuts.
- We get  $(SF) = .784 \pm .085$  (stat) with HenriB(Jan30) cuts and a modified method.
- Monte Carlo statistical uncertainty should improve shortly.
- Systematic on  $\boxed{F}^b/\boxed{F}^c$  is important – *not* completely accounted for with  $\sigma_{\boxed{F}^H}$ .
- We're studying systematics with a  $\chi^2$ -based fit method too –
  - Preliminary *Single tag*  $(SF)$  (with respect to truth)  $\sim .86 \pm .05$  with reasonable  $\boxed{F}^H$ .
  - (This is really pushing the small MC dataset we have)
  - Next: goodness-of-fit studies...